Feature

Widening the net on genome sequencing

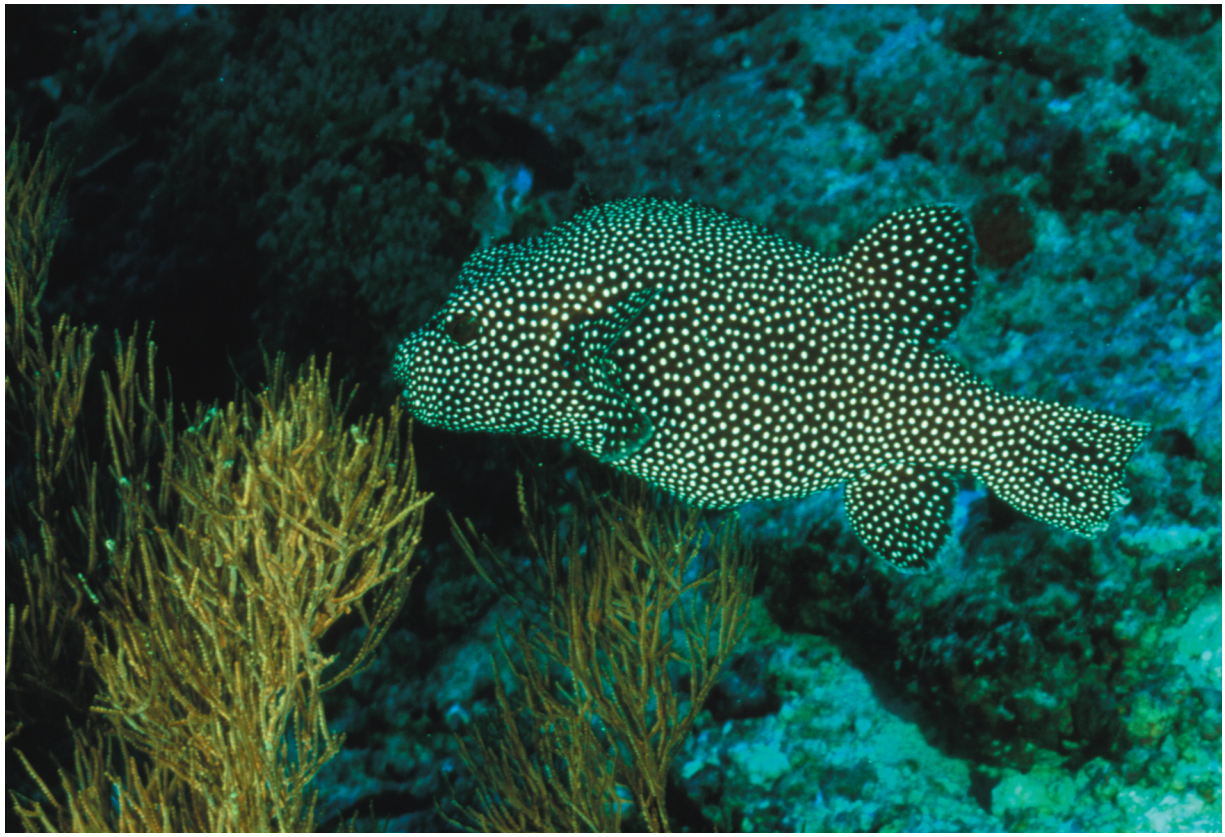
Genomes: With the DNA sequencing of most key model organisms finished or well on the way, one genomics institute is filling an important niche, using novel criteria to take on species from organisms that occupy a unique place in the evolutionary tree to others that have an ability to digest explosives. **Heather Dawes** reports.

At least for local scientists, it used to be that Walnut Creek, California was known mostly for its somewhat upscale shopping mall. But now this suburban town, located a half-hour's drive from Berkeley, is exporting more than Banana Republic couture – it also churns out a billion bases of raw DNA sequence data a month.

This is because Walnut Creek is home to the Production Genomics Facility, the sequencing powerhouse of the US Department of Energy's Joint Genome Institute. The JGI, which is funded and overseen by the DOE's Office of Science and administered by the University of California, was born in 1997, when the human genome sequencing efforts of three national labs – Los Alamos, Lawrence Livermore and Lawrence Berkeley – were consolidated to form a more efficient joint effort in the face of mounting competition from Celera Genomic's private sector advances.

Now, as the US government's largest sequencing center, the JGI is quickly becoming host to what might be called a third generation of genome projects: with most model organism genomes finished or well on the way, and the JGI's own work on the human genome just wrapping up, the institute is turning its attention to those non-model organisms with a special quality, whether it be a unique place in the evolutionary tree or the ability to digest explosives, that make them particularly interesting to government researchers and academic biologists alike.

How are the DOE's goals in genomics different from those of other US government agencies? According to JGI Interim Director Eddy Rubin, the DOE's interests in environmental aspects of biology mean that a wide spectrum of organisms are of potential interest for study, and that the DOE is in fact less limited than the NIH in this regard. 'In terms of 'what to sequence next?' we're much more interested in the broader biome;



Slimmed down: Puffer fish may be able to inflate themselves when under threat but at least one species has been shown to have a remarkably compact genome whose sequence is now nearing completion. (Photo: Science Photo Library.)

Stepping stone: Tunicates, as primitive chordates, occupy a crucial place in evolution and one species is attracting particular attention for those interested in the study of cis-acting regulatory elements. (Photo: Science Photo Library.)

we're dedicated to understanding basic biological questions — like carbon sequestration — with less relevance to medicine and more relevance to the basic biology of the planet.'

The model behind JGI's current endeavors, according to Rubin, is to take the rich set of resources the Institute has to offer in the way of sequencing power, bioinformatics resources, and functional genomics platforms, and make it available to communities of researchers to help them tackle their big biological questions. A prime example of such a collaboration is the role the JGI has played in the ongoing *Fugu rubripes* genome project. Fugu's remarkably compact genome – about 1/8th the size of our own – makes it potentially invaluable as a tool for comparative studies with mammalian genomes. At about the time the Fugu community, lead by Sydney Brenner, was looking for a group to take on the sequencing project, the JGI was looking for a new big project amenable to the whole-genome shotgun approach, and the match was made.

According to Dan Rokhsar, JGI's Associate Director for Computational Genomics, in addition to the JGI's substantial role in sequencing the Fugu genome the Institute also played the role of brokering another key collaboration: Singapore, eager to jump-start its national genomics program, agreed to bankroll additional Fugu sequencing from private sector players Celera and Myriad Genetics.

Sam Aparicio, of the Hutchinson/MRC Research Centre in Cambridge, UK and a leader on the Fugu project, says the JGI's part was critical. 'The role they've played has been extremely important,' says Aparicio, not only for catalyzing the start of the project, but for bringing the necessary assets together. 'It really wouldn't have gotten started without them.'

The Fugu project is now in the sequence finishing and annotation phase, with Singapore's Institute of Cell and Molecular Biology, the JGI, and other centers taking part, and an annotation jamboree planned for the coming year, following the publication of key data in *Science* last month.

With the Institute's sequencing prowess firmly established, it is now beginning to take a step further in its collaborative efforts, specifically in the direction of functional genomics. This relatively new endeavor has been developed substantially through the JGI's recent work on the tunicate *Ciona intestinalis.* A primitive chordate that is the subject of study for an international community of researchers, *Ciona* is of particular interest to those who want to understand development and the evolution of cis-acting regulatory elements, like UC Berkeley researcher Mike Levine. So, in another instance where the JGI's interests (here, branching out into functional genomics)

complemented those of the community, Levine and the Institute teamed up to establish the *Ciona* genome project, which has included a key collaboration with an established *Ciona* genomics group led by Nori Satoh in Japan.

As *Ciona* is an experimental organism, the functional genomics side of the project benefitted significantly from expertise Levine gained in his own lab, according to JGI's Dan Rokhsar. The Institute's successful efforts to scale up the identification of regulatory regions, which brought live animals and an additional wet lab component to the JGI, has identified 20–30 new enhancers so far. Annotation of the genome sequence is also ongoing, boosted by a meeting hosted by the JGI this spring that drew an international group of tunicate researchers.

What is the Institute taking on next? Three kingdoms, as it turns out: genome projects for *Chlamydomonas*, *Xenopus tropicalis*, and Poplar are all in the works. Both *Chlamydomonas* and Poplar are of special interest to the DOE for their genetic tractability and potential use in understanding carbon sequestration. Furthest along is the *Chlamydomonas* project, with sequencing wrapping up soon and an annotation meeting planned for the fall.

The *Xenopus* project is a joint effort with academic researchers led by Rob Grainger of the University of Virginia and Richard Harland of UC Berkeley, and is partially funded by the EPA, owing to the frog's utility as a sentinel organism for environmental toxicity. A genetically amenable cousin of *X. laevis, tropicalis* has a much smaller, diploid genome, and a growing community of researchers developing molecular and genetic tools designed for studying it as a key model organism.

According to Paul Richardson of the JGI's functional genomics group, all three projects got started through close discussions with the various research communities involved, a process that JGI hopes to see mature into a more established mechanism involving peer-reviewed proposals and hands-on collaborations with lab heads and postdocs.

In the coming years, the JGI's resources will be split fairly evenly between DOE initiatives and projects based in the broader scientific community, according to Marvin Frazier, Director of the Office of Science's Life Sciences division, which oversees the JGI and related projects. In the longer term, a good portion of the JGI's activity will likely stem from the DOE's newly unveiled Genomes to Life initiative, a wide-ranging, long-term plan that will employ the resources of the Office of Biological and Environmental Science, as well as the Office of Advanced Scientific Computing, to understand at an integrated, 'systems' level the biology underlying energy production, carbon management, and bioremediation.

With those goals, the initiative should give the JGI an opportunity to shine in an area it knows well: microbial genomics. With much experience gained through its past involvement with the DOE's Microbial Genomes Project, and a multitude of ongoing projects, the JGI will be called upon to tackle the genomes not only of single organisms with unusual metabolic characteristics, but also so-called microbial consortia – complex communities of microbes, often not culturable in the laboratory, that are isolated from harsh or toxic environments.

The JGI has their work cut out for them. Given Genomes to Life's stated objective of contributing to 'a fundamental, comprehensive, and systematic understanding of life,' calling the initiative ambitious borders on understatement. This is something readily acknowledged by the DOE's Marvin Frazier: 'It's audacious! But so was the Human Genome Project when we started.'

Following the JGI's successful contribution to that groundbreaking enterprise, the Institute now seems poised to once again help make the audacious dreams of researchers a reality.

Quick guide

Opsins and melanopsins

Russell G. Foster and James Bellingham

What are opsins? Opsins are generally considered members of the superfamily of G-protein coupled receptors. But not all opsins activate a G-protein. Their distinguishing features are a 7 transmembrane α-helical structure, and an ability to bind a vitamin A chromophore, retinaldehyde, using a lysine in the 7th α-helix. The range of amino acid identity among opsin families is 22–40%, and they have ~17% identity to other hepta-helical receptor families.

How do they work? Two well established functions are those of photosensor and photoisomerase (Figure 1). In the photosensory opsins the chromophore, 11-*cis*-retinal, is located in a cage formed by the α-helices. Light is absorbed by 11-*cis*-retinal which is photoisomerised to all-*trans*. This conformational change allows the opsin to bind the α-subunit of the G-protein transducin. In the rods and cones of the retina G-protein activation leads to hyper-polarisation of the photoreceptor through cGMP-gated cation channels. The photoisomerase function is exhibited by retinal G-protein coupled receptor (RGR) opsin, expressed in the retinal pigment epithelium. It binds all-*trans*-retinal and uses light to convert it to the 11-*cis* configuration. Photoisomerisation is not thought to activate a G-protein but supplies the rod and cone opsins with 11-*cis* chromophore.

How many opsin families are there in vertebrates? On the basis of sequence similarity there are 14: 7 are photosensory, including the 4 cone and 1 rod-opsin families, pineal-opsins, and vertebrate ancient (VA) opsins,